DOCUMENT RESUME

ED 258 143 CS 008 051

AUTHOR Olson, Arthur V.
TITLE Readability Formulas--Fact or Fiction.
PUB DATE [84]
NOTE 16p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Content Analysis; Elementary Secondary Education; *Readability; *Readability Formulas; Reading Instruction; *Reading Materials; Reading Material Selection; Reading Research

ABSTRACT

Four readability formulas were analyzed to determine the reading grade equivalence of instructional materials at the middle grades and above: the Dale-Chall formula, the Gunning-Fog Index, the Flesch Reading Ease Formula, and the McLaughlin SMOG Grading. In addition, Spache and Wheeler/Smith formulas were analyzed for evaluating primary grade materials. These analyses showed some problems with the readability formulas, including the following: (1) the formulas ignore much of the current knowledge and theoretical bases of the reading process, (2) they rest on a shaky statistical base, and (3) there is considerable variation in grade level designation for the reading selections. Furthermore, if the readability formulas are based upon the McCall-Crabbs Standard Test Lessons as their criterion for a standard, then the formulas themselves should indicate a difference in the difficulty levels from Books A-D, which they do not to the extent assumed by those who used the lessons as the criterion for the development of their formulas. The conclusions, based upon the findings, are that the readability formulas were based upon unsound criteria, that the correlation research to support the use of other formulas (strength by association) has similar weaknesses, and that the folklore extrapolated from the unfounded assumptions places the question of the use of readability formulas in great jeopardy. (EL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED258143

Readability Formulas- Fact or Fiction

Dr. Arthur V. Olson, Professor

Chairman, Dept. of Communication and

Social Foundations

Faculty of Education

University of Victoria

Victoria, B.C. - V8W 2Y2

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Running Head: READABILITY FORMULAS

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Arthur Olson

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

Readability Formulas--Fact or Fiction

The prospect of measuring the readability of a text by means of a simple formula has been a reality since 1923 with more than fifty having been published to date. Application of the formulas has an extensive history in the writing of textbooks, for all content areas inclusive of basal readers, selection of instructional materials for specific target populations and in the construction of achievement tests. In several states within the United States and in provinces in Canada, materials are often selected by the assigned government agency for approved use only if they fit specified readability criteria.

The readability formulas and the scores (reading grade equivalents) resulting from their application are indicated with increasing frequency on brochures and catalogues as a means of increasing selling appeal. Materials that are written for multilevel use and as remedial material are certainly more marketable if the formula and scores are reported that provide information on the grade range.

Most of the formulas in current use are based upon the common factors of vocabulary, measured by syllable count, and of syntax, measured by average sentence length. Four commonly used readability formulas for the middle grades and above were analyzed in this study: (a) Dale-Chall, (b) Gunning-Fox Index, (c) Flesch Reading Ease Formula, and (d) McLaughlin SMOG Grading. In addition, Spache and Wheeler/Smith formulas were analyzed for evaluating primary grade materials.

The first four of the readability formulas share a common problem in their development with respect to the criterion. The McCall-Crabbs Standard Test Lessons in Reading (McCall & Crabbs, 1925, 1950, 1961) were never standardized, nor were they intended for the research purposes for which they have been employed. Stevens (1980), in a personal communication in 1978 with William McCall, reported that McCall said, "When, last year, I learned for the first time the number of readability formulae resting on my G (grade level) scores, I was vastly surprised. . . . Probably all the formulae were defensible during the crude early days of scientific education. The formulae builders never approached me as you have done." In spite of this denunciation of the use of the McCall-Crabbs "graded" selections, earlier comments by some readability authorities have been widely read and accepted. Klare (1974-75) for example stated: "These lessons (the McCall-Crabbs Standard Test Lessons) have been convenient statistically because they are a large number of reading passages, covering a wide range of difficulty, resting upon extensive testing and providing detailed grade scores."

A predictive formula for readability can be only as good as the criterion upon which it is based. The basic assumption is that the McCall-Crabbs material represents a valid and reliable criterion. In order to understand better the magnitude of the problem, a closer look at the formulas used for this study may be in order.

The Flesch Reading Ease Formula is a relatively simple formula

with factors measuring aspects of the difficulty of reading single words and sentences. Numerous validity studies for this formula have been reported (Klare, 1963). These studies use a variety of different criteria, such as comprehension test scores and expert judgment. The highest correlations reported are usually between the Flesch and other readability formulas. These data are not particularly useful since commonsense alone would indicate that the correlation would be high if the internal factors for measuring readability and the criterion base were the same. The results of the correlation on the criterion measure, McCall-Crabbs Lessons, with all formulas have a measure of correlation at a significant level.

The sentence and word factors in the Dale-Chall Formula (1948) are the same as in the Flesch formula. It can be argued that the percentage of words outside the Dale list of 3,000 (X_1 or Dale score) measures the same influence as the Flesch syllable count. The word factor may be seen as a measure of semantic difficulty since the syllable count is a reflection of word length and higher syllable counts reflect a measure of degree of usage, and therefore difficulty of meaning. Words found in the Dale list were selected by questioning a sample of fourth-graders who checked the words they knew. A word was included on the list when 80% of the sample had sight knowledge of the words. Other word lists (Thorndike & Lorge, 1944; Carroll, Davies & Richman, 1971) were compiled on the basis of the number of times a lexical item appeared in a cross-genre sample of selected texts.

The Gunning Fog Index (1952) is clearly similar to that of the Flesch and Dale-Chall in that the factors remain the same but the counting of three-syllable words is easier and requires less time than the syllable count required by Flesch.

The SMOG (McLaughlin, 1969) predicts readability on the basis of one factor, polysyllabic words. It is therefore limited to those factors related to the identification and processing of words.

A comparison of the Fog Index and SMOG grading showed the highest correlation (.67). A comparison of readability grade levels with Dale-Chall showed that the SMOG grades were usually two grades higher. This higher grade level designation is probably due to the use of a more severe criterion for the SMOG grading than the Dale-Chall formula: McLaughlin used one hundred percent success on McCall-Crabbs' materials as the criterion for comprehension, while Dale and Chall employed a fifty percent comprehension score as their criterion.

The Spache formula (1953) was designed for use with primary grade materials, as was the Wheeler and Smith formula. The sample (224) was drawn from 152 primary-grade textbooks according to level of classroom use based upon publishers' designations. The Spache is the most highly correlated with the Dale-Chall (.63) which is not surprising since the bases for the formulas have a great deal in common.

The Wheeler and Smith formula (1954) resembles the Spache in the sense that the criterion used was determined by publishers' grade designations for books. In the case of the Wheeler and Smith formula,

combined means of nine basic reading series were used. Wheeler and Smith, who took a random word-count of the nine reading series, found the mean percentage of polysyllabic words for each level and the mean unit length for each level. The combined means were established as the criterion for each level. They felt that they might be criticized for giving equal weight to unit length, which is similar to sentence length. They were concerned that too much weight might be given to a factor generally considered to be less important than vocabulary. To counter this criticism, they pointed out that in every case throughout the nine series there was a clear increase in sentence length from grade to grade.

The application of readability formulas raises several concerns that should be mentioned. First, the formulas ignore much of the current knowledge and theoretical bases of the reading process. Most authorities would agree that the goal of reading is the comprehension of the written message. Because most of the formulas are based upon sentence length and word complexity (based on the number of syllables in the word) other factors that make a text difficult are measured only indirectly, if at all. The degree of abstraction, complexity and relationship of ideas, rhetorical structure, degree of discourse cohesion, dialect, text organization, and conceptual background of the reader are all eliminated from consideration in the formula approach. In addition, use to which the reader is to put the material is another missing factor in readability formulas. They do not take into account

the factors of motivation, values, interest or purpose.

The second factor, which is the main focus of this paper, is the shaky statistical base upon which the formulas rest. A look at the pertinent assumptions upon which readability formulas rest may help in drawing some conclusions with respect to this weakness. The primary question relates to the criterion for the formulas.

Four of the six formulas used the McCall-Crabbs Standard Test Lessons in Reading as the independent criterion of difficulty to establish validity. The remainder of the formulas often cite their correlation with the first four in order to determine validity. Although this is a questionable technique, it is frequently used to establish a relationship by association. The McCall-Crabbs Standard Test Lessons in Reading consist of four booklets (A-D), each containing approximately seventy graded passages with multiple choice questions. The grade level assigned to each is alleged to be based on the number of correct responses by pupils of known reading achievement with each of the passages arranged from least difficult to most difficult within each booklet, and each book containing more difficult material in order.

If the McCall-Crabbs lessons are to be used as the basis to measure readability then there should be proof that the lessons increase in difficulty (reading grade equivalence) from the beginning of any given booklet to the end, and from Book A to Book D. This should certainly be true if the readability formulas used to determine the

difficulty level were the formulas that employed the McCall-Crabbs Lessons as the criterion of difficulty.

If we consider (a) the number of sentences, and (b) the syllables per one hundred words as a measure of difficulty of material, the averages for these two factors do not support the assumption that Books A-D increase in complexity of material. The average number of sentences per one hundred words is: 7.0, 7.4, 7.0 and 7.5 for Books A-D respectively. The average number of syllables per one hundred words is: 124.3, 121.7, 123.8 and 131.4 for Books A-D respectively. These data do not support the assumption that the McCall-Crabbs Test Lessons in Reading provide graded reading selections increasing in difficulty.

In comparing the mean grade level for Books A-D using the Dale-Chall formula, it is found that there is a statistically significant difference (.01 level) between the first third and the last third of reading selections in Books A and D. The mean grade level for each third of the books, however, appears to be inappropriate for the grade level selections as assigned by McCall and Crabb.

The mean grade level for Books A-D, using the Fog Index, indicates a statistically significant difference (.01 level) between the first third and last third of reading selections in Books B and C. Within each book (B and C) the last third indicates considerably higher grade level than that designated by McCall-Crabb for the range of reading ability.

The mean grade level for Books B, C and D, using the Flesch formula,

indicates a statistically significant difference (.01 level) between the first third and last third of the reading selections. The differences between the means for Book B were less than one year, while Books C and D indicate mean differences slightly in excess of two years.

The mean grade level for Books A-D, using the SMOG Index, indicates no statistically significant difference (.01 level) between the first third and last third of the reading selections in any of the McCall-Crabbs books.

The mean grade level for Book C, using the Spache Index, indicates a statistically significant difference (.01 level) between the first third and the last third of the reading selections. Analysis of the means reflects one grade level difference.

The mean grade level for Books A and C, using the Wheeler-Smith Formula, indicates a statistically significant difference (.01 level) between the first third and the last third of the reading selections. The difference between the grade level means for Book A was less than one year.

Among the formulas there is considerable variation in grade level designation for the reading selections, both within each level book and among the books of the McCall-Crabbs Standard Reading Lessons, even in light of the knowledge that the Spache Index and the Wheeler-Smith Formula are for use in primary grade material. The results only increase concern about using any formula that uses the McCall-Crabbs Standard Test Lessons as a criterion for development, or that bases its validity

upon the correlation with other formulas as its justification for application. An added concern that also enters into the discussion is the error factor of ± 1.0 year applied to any grade level indicated by a formula.

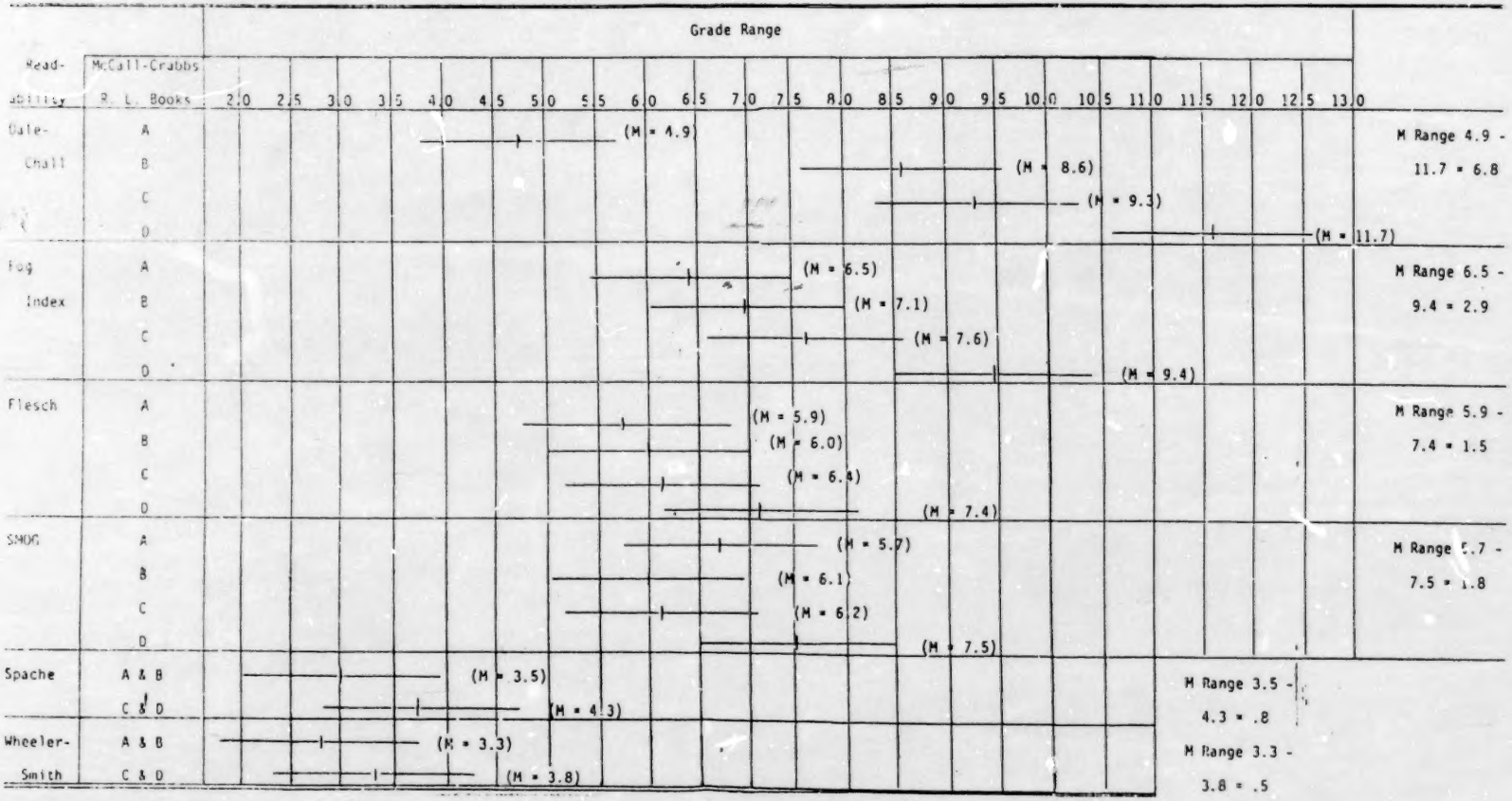
Another interpretation of the reading levels of the McCall-Crabbs Lessons based upon the formulas cited is reflected in Table I. The Dale-Chall formula appears to indicate a wider range of reading levels than any of the other formulas employed, with a mean grade level range of 6.8 years. There is an obvious clustering of similar level reading selections in Books B and C with a mean difference of only seven months. The mean grade level range for the Books A-D, using the Fog Index, is 2.9 years with a mean grade level range of 1.5 and 1.8 for the Fog and the Flesch Readability formulas respectively. The Spache and the Wheeler-Smith Readability formulas indicate a much lower reading grade level for Books A and B than the other formulas cited, and the range between Books A-B, C-D is extremely small: .8 months and .5 months.

Conclusion

Chall, in a recent publication (1984) indicated that readability formulas predict difficulty in terms of reading grade equivalents and not of grade placement. Chall assumes that a given readability grade level means that a student who reads at the same given grade level on a standardized reading test can be expected to read and understand material at about seventy-five percent accuracy on a multiple choice test, and that the level is reflected by the readability formulas.

Table 1

Analysis of Readability Level of the McCall-Crabbs Standard Reading Lessons



Note: The range reflects the mean reading level for each book ± 1.0 (error).

BEST COPY AVAILABLE

The data provided by this report do not support Chall's assumptions, nor those made by others who would have us use readability formulas to determine the reading grade equivalence of instructional materials. If the readability formulas are based upon the McCall-Crabbs Standard Test Lessons as their criterion for a standard, then the formulas themselves should indicate a difference in the difficulty levels from Books A-D. It is obvious from the data that the McCall-Crabbs Standard Test Reading Lessons do not increase in reading difficulty, either within the individual books or among the books, to the extent assumed by those who used the lessons as the criterion for the development of their formulas. The basic assumptions upon which readability formulas have been developed are indeed suspect and the conclusions and extended logic that have developed with respect to their current, and previous, use are erroneous. The only reasonable conclusions that can be reached, based upon the findings, are that the readability formulas were based upon unsound criteria, that the correlation research to support the use of other formulas (strength by association) has similar weaknesses, and finally that the folklore extrapolated from the unfounded assumptions places the question of the use of readability formulas in great jeopardy.

References

Carroll, J. B., Davies, P., & Richman, B. (1971). The American heritage word frequency book. Boston, MA: Houghton Mifflin.

Chall, J. (1984). Readability and prose comprehension: Continuities and discontinuities. In J. Flood (Ed.), Understanding reading comprehension (pp. 233-264). Newark, DE: International Reading Association.

Dale, E., & Chall, J. (1948, January/February). Formula for predicting readability. Educational Research Bulletin, 27, 11-20, 28, 37-54.

Flesch, R. (1948, June). New readability yardstick. Journal of Applied Psychology, 32, 221-223.

Gunning, R. (1968, Winter). The Fog Index after twenty years. The Journal of Business Communication, 6, 3-13.

Klare, G. R. (1963). The measurement of readability. Ames, IA: Iowa State University Press.

Klare, G. R. (1974-75). Assessing readability. Reading Research Quarterly, 10(1), 62-102.

McCall, W. A., & Crabbs, L. M. (1925, 1950, 1961). Standard test lessons in reading. New York: Teachers College Press.

Spache, G. (1953, March). New readability formula for primary grade reading materials. Elementary School Journal, 53, 410-463.

Stevens, K. C. (1980, January). Readability formulae and McCall-Crabbs standard test lessons in reading. The Reading Teacher, 410-415.

Thorndike, E. L., & Lorge, I. (1944). The teacher's wordbook of 30,000 words. New York: Bureau of Publications, Columbia University, Teachers College.

Wheeler, L. R., & Smith, E. H. (1954). A practical readability formula for the classroom teacher in the primary grades. Elementary English, 31, 397-399.